

11



PCT10

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:54

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

ENTERED

3 <110> APPLICANT: Luet, Wong
 4 Jonathan, Jones
 6 <120> TITLE OF INVENTION: PROCESS FOR OXIDISING AROMATIC COMPOUNDS
 8 <130> FILE REFERENCE: P02353US1 / 10112404 / N.76277B
 10 <140> CURRENT APPLICATION NUMBER: US 10/018,730A
 C--> 11 <141> CURRENT FILING DATE: 2002-04-04
 13 <150> PRIOR APPLICATION NUMBER: GB 9914373.7
 14 <151> PRIOR FILING DATE: 1999-06-18
 16 <150> PRIOR APPLICATION NUMBER: PCT/GB00/02379
 17 <151> PRIOR FILING DATE: 2000-06-19
 19 <160> NUMBER OF SEQ ID NOS: 18
 21 <170> SOFTWARE: PatentIn version 3.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 1242
 25 <212> TYPE: DNA
 26 <213> ORGANISM: Pseudomonas putida
 28 <400> SEQUENCE: 1
 29 acgactgaaa ccatacaaaag caacgcgaat cttgcccctc tgccacccca tgtgccagag 60
 31 cacctggtat tcgacttcga catgtacaat ccgtcgaaac tgtctgccgg cgtgcaggag 120
 33 gcctgggcag ttctgcaaga atcaaacgta ccggtctggt tgtggactcg ctgcaacggc 180
 35 ggacactgga tcgccactcg cggccaactg atccgtgagg cctatgaaga ttaccgccac 240
 37 ttttccagcg agtgcccgtt catccctcgt gaagccggcg aagcctacga cttcattccc 300
 39 acctcgatgg atccgcccga gcagcgccag tttcgtgcgc tggccaacca agtggttggc 360
 41 atgccggtgg tggataagct ggagaaccgg atccaggagc tggcctgctc gctgatcgag 420
 43 agcctgcgcc cgcaaggaca gtgcaacttc accgaggact acgccgaacc cttcccgcata 480
 45 cgcactttca tgctgctcgc aggtctaccg gaagaagata tccgcactt gaaataccta 540
 47 acggatcaga tgaccgcgtc ggatggcagc atgaccttcg cagaggccaa ggaggcgtc 600
 49 tacgactatc tgataccgat catcgagcaa cgcaggcaga agccgggaac cgacgctatc 660
 51 agcatcggtg ccaacggcca ggtcaatggg cgaccgatca ccagtgcga agccaagagg 720
 53 atgtgtggcc tgttactggt cggcggcctg gatacgggtg tcaatttcct cagcttcagc 780
 55 atggagtccc tggccaaaag cccggagcat cgccaggagc tgatcgagcg tcccgcagct 840
 57 attccagccg cttgcgagga actactccgg cgcttctcgc tggttgccga tggccgcac 900
 59 ctcacctccg attacgagtt tcatggcggt caactgaaga aaggtgacca gatcctgcta 960
 61 ccgcagatgc tgtctggcct ggatgagcgc gaaaacgcct gcccgatgca cgtcgacttc 1020
 63 agtcgccaaa aggtttcaca caccaccttt ggccacggca gccatctgtg ccttgccag 1080
 65 cacctggccc gccgggaaat catcgtcacc ctcaagggaat ggctgaccag gattcctgac 1140
 67 ttctccattg ccccggtgac ccagattcag cacaagagcg gcacgtcag cggcgtgcag 1200
 69 gcactccctc tggctctgga tccggcgact accaaagcgg ta 1242
 72 <210> SEQ ID NO: 2
 73 <211> LENGTH: 414
 74 <212> TYPE: PRT
 75 <213> ORGANISM: Pseudomonas putida
 77 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:54

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

```

79 Thr Thr Glu Thr Ile Gln Ser Asn Ala Asn Leu Ala Pro Leu Pro Pro
80 1 5 10 15
83 His Val Pro Glu His Leu Val Phe Asp Phe Asp Met Tyr Asn Pro Ser
84 20 25 30
87 Asn Leu Ser Ala Gly Val Gln Glu Ala Trp Ala Val Leu Gln Glu Ser
88 35 40 45
91 Asn Val Pro Asp Leu Val Trp Thr Arg Cys Asn Gly Gly His Trp Ile
92 50 55 60
95 Ala Thr Arg Gly Gln Leu Ile Arg Glu Ala Tyr Glu Asp Tyr Arg His
96 65 70 75 80
99 Phe Ser Ser Glu Cys Pro Phe Ile Pro Arg Glu Ala Gly Glu Ala Tyr
100 85 90 95
103 Asp Phe Ile Pro Thr Ser Met Asp Pro Pro Glu Gln Arg Gln Phe Arg
104 100 105 110
107 Ala Leu Ala Asn Gln Val Val Gly Met Pro Val Val Asp Lys Leu Glu
108 115 120 125
111 Asn Arg Ile Gln Glu Leu Ala Cys Ser Leu Ile Glu Ser Leu Arg Pro
112 130 135 140
115 Gln Gly Gln Cys Asn Phe Thr Glu Asp Tyr Ala Glu Pro Phe Pro Ile
116 145 150 155 160
119 Arg Ile Phe Met Leu Leu Ala Gly Leu Pro Glu Glu Asp Ile Pro His
120 165 170 175
123 Leu Lys Tyr Leu Thr Asp Gln Met Thr Arg Pro Asp Gly Ser Met Thr
124 180 185 190
127 Phe Ala Glu Ala Lys Glu Ala Leu Tyr Asp Tyr Leu Ile Pro Ile Ile
128 195 200 205
131 Glu Gln Arg Arg Gln Lys Pro Gly Thr Asp Ala Ile Ser Ile Val Ala
132 210 215 220
135 Asn Gly Gln Val Asn Gly Arg Pro Ile Thr Ser Asp Glu Ala Lys Arg
136 225 230 235 240
139 Met Cys Gly Leu Leu Leu Val Gly Gly Leu Asp Thr Val Val Asn Phe
140 245 250 255
143 Leu Ser Phe Ser Met Glu Phe Leu Ala Lys Ser Pro Glu His Arg Gln
144 260 265 270
147 Glu Leu Ile Glu Arg Pro Glu Arg Ile Pro Ala Ala Cys Glu Glu Leu
148 275 280 285
151 Leu Arg Arg Phe Ser Leu Val Ala Asp Gly Arg Ile Leu Thr Ser Asp
152 290 295 300
155 Tyr Glu Phe His Gly Val Gln Leu Lys Lys Gly Asp Gln Ile Leu Leu
156 305 310 315 320
159 Pro Gln Met Leu Ser Gly Leu Asp Glu Arg Glu Asn Ala Cys Pro Met
160 325 330 335
163 His Val Asp Phe Ser Arg Gln Lys Val Ser His Thr Thr Phe Gly His
164 340 345 350
167 Gly Ser His Leu Cys Leu Gly Gln His Leu Ala Arg Arg Glu Ile Ile
168 355 360 365
171 Val Thr Leu Lys Glu Trp Leu Thr Arg Ile Pro Asp Phe Ser Ile Ala
172 370 375 380
175 Pro Gly Ala Gln Ile Gln His Lys Ser Gly Ile Val Ser Gly Val Gln

```

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:54

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

```

176 385          390          395          400
179 Ala Leu Pro Leu Val Trp Asp Pro Ala Thr Thr Lys Ala Val
180          405          410
183 <210> SEQ ID NO: 3
184 <211> LENGTH: 3150
185 <212> TYPE: DNA
186 <213> ORGANISM: Bacillus megaterium
188 <400> SEQUENCE: 3
189 atgacaatta aagaaatgcc tcagccaaaa acgtttggag agcttaaaaa tttaccgtta 60
191 ttaaaccacag ataaaccggt tcaagctttg atgaaaattg cggatgaatt aggagaaatc 120
193 tttaaattcg aggcgcctgg tcgtgtaacg cgctacttat caagtcagcg tctaattaaa 180
195 gaagcatgcg atgaatcacg ctttgataaa aacttaagtc aagcgcctaa atttgtacgt 240
197 gattttgcag gagacgggtt atttacaagc tggacgcagt aaaaaaattg gaaaaaagcg 300
199 cataatatct tacttccaag cttcagtcag caggcaatga aaggctatca tgcgatgatg 360
201 gtcgatatcg ccgtgcagct tgttcaaaaag tgggagcgct taaatgcaga tgagcatatt 420
203 gaagtaccgg aagacatgac acgtttaacg cttgatacaa ttggtctttg cggctttaac 480
205 tategcttta acagctttta ccgagatcag cctcatccat ttattacaag tatggtccgt 540
207 gcactggatg aagcaatgaa caagctgcag cgagcaaatc cagacgaccc agcttatgat 600
209 gaaaacaagc gccagtttca agaagatata aaggtgatga acgacctagt agataaaatt 660
211 attgcagatc gcaaagcaag cgggtgaaca agcgatgatt tattaacgca tatgctaaac 720
213 ggaaaagatc cagaaacggg tgagccgctt gatgacgaga acattcgcta tcaaattatt 780
215 acattcttaa ttgcgggaca cgaacaacaa agtggctttt tatcatttgc gctgtatttc 840
217 ttagtgaaaa atccacatgt attacaaaaa gcagcagaag aagcagcacg agttctagta 900
219 gatcctgctc caagctacaa acaagtcaaa cagcttaaat atgtcggcat ggtcttaaac 960
221 gaagcgcctg ctttatggcc aactgctcct gcgttttccc tatatgcaa aagaagatacg 1020
223 gtgcttgagg gagaatatcc tttagaaaaa ggcgacgaac taatggttct gattcctcag 1080
225 cttcaccgtg ataaaaaat ttggggagac gatgtggaag agttccgtcc agagcgtttt 1140
227 gaaaatccaa gtgcgattcc gcagcatgcg tttaaaccgt ttggaaacgg tcagcgtgcg 1200
229 tgtatcggtc agcagttcgc tcttcatgaa gcaacgctgg tacttggtat gatgctaaaa 1260
231 cactttgact ttgaagatca tacaactac gagctggata ttaaagaaac tttaacgtta 1320
233 aaacctgaag gctttgtggt aaaagcaaaa tcgaaaaaaa ttccgcttg ggtattcctt 1380
235 tcacctagca ctgaacagtc tgccaaaaaa gcacgcaaaa aggcagaaaa cgctcataat 1440
237 acgccgctgc ttgtgctata cggttcaaat atgggaacag ctgaaggaaac ggcgcgatg 1500
239 ttagcagata ttgcaatgag caaaggattt gcaccgcagg tcgcaacgct tgattcacac 1560
241 gccggaaaatc ttccgcgcga aggagctgta ttaattgtaa cggcgcttta taacggtcat 1620
243 ccgcctgata acgcaaagca atttgtcgac tggttagacc aagcgtctgc tgatgaagta 1680
245 aaaggcgctt gctactccgt atttggatgc ggcgataaaa actgggctac tacgtatcaa 1740
247 aaagtgcctg cttttatcga tgaacgcctt gccgctaaag gggcagaaaa catcgtgac 1800
249 cgcggtgaag cagatgcaag cgacgacttt gaaggcacat atgaagaatg gcgtgaacat 1860
251 atgtggagtg acgtagcagc ctactttaac ctgcacattg aaaacagtga agataataaa 1920
253 tctactcttt cacttcaatt tgcgcacagc gccgcggata tgccgcttgc gaaaatgcac 1980
255 ggtgcgtttt caacgaacgt cgtagcaagc aaagaacttc aacagccagg cagtgcacga 2040
257 agcacgcgac atcttgaat tgaacttcca aaagaagctt cttatcaaga aggagatcat 2100
259 ttaggtgtta ttccctgcaa ctatgaagga atagtaaacc gtgtaacagc aaggttcggc 2160
261 ctagatgcat cacagcaaat ccgtctggaa gcagaagaag aaaaattagc tcatttgcca 2220
263 ctgcataaaa cagtatccgt agaagagctt ctgcaatacg tggagcttca agatcctgtt 2280
265 acgcgcacgc agcttcgcgc aatggctgct aaaacggtct gcccgccgca taaagtagag 2340
267 cttgaagcct tgcttgaaaa gcaagcctac aaagaacaag tgctggcaaa acgtttaaca 2400
269 atgcttgaac tgcttgaaaa ataccggcg tgtgaaatga aattcagcga atttatcgcc 2460

```

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:54

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

```

271 cttctgccaa gcatacgccc gcgctattac tcgatttctt catcacctcg tgtcgatgaa 2520
273 aaacaagcaa gcatcacggt cagcgttgtc tcaggagaag cgtggagcgg atatggagaa 2580
275 tataaaggaa ttgcgtcgaa ctatcttgcc gagctgcaag aaggagatac gattacgtgc 2640
277 tttatttcca caccgcagtc agaatttacg ctgccaaaag accctgaaac gccgcttatac 2700
279 atggtcggac cgggaacagg cgtcgcgcgcg tttagaggct ttgtgcaggc gcgcaaacag 2760
281 ctaaaagaac aaggacagtc acttgagaaa gcacatttat acttcggctg ccgttcacct 2820
283 catgaagact atctgtatca agaagagctt gaaaacgccc aaagcgaagg catcattacg 2880
285 cttcataccg cttttttctg catgccaaat cagccgaaaa catacgttca gcacgtaatg 2940
287 gaacaagacg gcaagaaatt gattgaactt cttgatcaag gagcgcactt ctatatttgc 3000
289 ggagacggaa gccaaatggc acctgccgtt gaagcaacgc ttatgaaaag ctatgctgac 3060
291 gttcaccaag tgagtgaagc agacgctcgc ttatggctgc agcagctaga agaaaaaggc 3120
293 cgatacgcaa aagacgtgtg ggctgggtaa 3150
296 <210> SEQ ID NO: 4
297 <211> LENGTH: 1049
298 <212> TYPE: PRT
299 <213> ORGANISM: Bacillus megaterium
301 <400> SEQUENCE: 4
303 Met Thr Ile Lys Glu Met Pro Gln Pro Lys Thr Phe Gly Glu Leu Lys
304 1 5 10 15
307 Asn Leu Pro Leu Asn Thr Asp Lys Pro Val Gln Ala Leu Met Lys
308 20 25 30
311 Ile Ala Asp Glu Leu Gly Glu Ile Phe Lys Phe Glu Ala Pro Gly Arg
312 35 40 45
315 Val Thr Arg Tyr Leu Ser Ser Gln Arg Leu Ile Lys Glu Ala Cys Asp
316 50 55 60
319 Glu Ser Arg Phe Asp Lys Asn Leu Ser Gln Ala Leu Lys Phe Val Arg
320 65 70 75 80
323 Asp Phe Ala Gly Asp Gly Leu Phe Thr Ser Trp Thr His Glu Lys Asn
324 85 90 95
327 Trp Lys Lys Ala His Asn Ile Leu Leu Pro Ser Phe Ser Gln Gln Ala
328 100 105 110
331 Met Lys Gly Tyr His Ala Met Met Val Asp Ile Ala Val Gln Leu Val
332 115 120 125
335 Gln Lys Trp Glu Arg Leu Asn Ala Asp Glu His Ile Glu Val Pro Glu
336 130 135 140
339 Asp Met Thr Arg Leu Thr Leu Asp Thr Ile Gly Leu Cys Gly Phe Asn
340 145 150 155 160
343 Tyr Arg Phe Asn Ser Phe Tyr Arg Asp Gln Pro His Pro Phe Ile Thr
344 165 170 175
347 Ser Met Val Arg Ala Leu Asp Glu Ala Met Asn Lys Leu Gln Arg Ala
348 180 185 190
351 Asn Pro Asp Asp Pro Ala Tyr Asp Glu Asn Lys Arg Gln Phe Gln Glu
352 195 200 205
355 Asp Ile Lys Val Met Asn Asp Leu Val Asp Lys Ile Ile Ala Asp Arg
356 210 215 220
359 Lys Ala Ser Gly Glu Gln Ser Asp Asp Leu Leu Thr His Met Leu Asn
360 225 230 235 240
363 Gly Lys Asp Pro Glu Thr Gly Glu Pro Leu Asp Asp Glu Asn Ile Arg
364 245 250 255

```

RAW SEQUENCE LISTING

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:54

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

```

367 Tyr Gln Ile Ile Thr Phe Leu Ile Ala Gly His Glu Thr Thr Ser Gly
368          260          265          270
371 Leu Leu Ser Phe Ala Leu Tyr Phe Leu Val Lys Asn Pro His Val Leu
372          275          280          285
375 Gln Lys Ala Ala Glu Glu Ala Ala Arg Val Leu Val Asp Pro Ala Pro
376          290          295          300
379 Ser Tyr Lys Gln Val Lys Gln Leu Lys Tyr Val Gly Met Val Leu Asn
380 305          310          315          320
383 Glu Ala Leu Arg Leu Trp Pro Thr Ala Pro Ala Phe Ser Leu Tyr Ala
384          325          330          335
387 Lys Glu Asp Thr Val Leu Gly Gly Glu Tyr Pro Leu Glu Lys Gly Asp
388          340          345          350
391 Glu Leu Met Val Leu Ile Pro Gln Leu His Arg Asp Lys Thr Ile Trp
392          355          360          365
395 Gly Asp Asp Val Glu Glu Phe Arg Pro Glu Arg Phe Glu Asn Pro Ser
396          370          375          380
399 Ala Ile Pro Gln His Ala Phe Lys Pro Phe Gly Asn Gly Gln Arg Ala
400 385          390          395          400
403 Cys Ile Gly Gln Gln Phe Ala Leu His Glu Ala Thr Leu Val Leu Gly
404          405          410          415
407 Met Met Leu Lys His Phe Asp Phe Glu Asp His Thr Asn Tyr Glu Leu
408          420          425          430
411 Asp Ile Lys Glu Thr Leu Thr Leu Lys Pro Glu Gly Phe Val Val Lys
412          435          440          445
415 Ala Lys Ser Lys Lys Ile Pro Leu Gly Gly Ile Pro Ser Pro Ser Thr
416          450          455          460
419 Glu Gln Ser Ala Lys Lys Ala Arg Lys Lys Ala Glu Asn Ala His Asn
420 465          470          475          480
423 Thr Pro Leu Leu Val Leu Tyr Gly Ser Asn Met Gly Thr Ala Glu Gly
424          485          490          495
427 Thr Ala Arg Asp Leu Ala Asp Ile Ala Met Ser Lys Gly Phe Ala Pro
428          500          505          510
431 Gln Val Ala Thr Leu Asp Ser His Ala Gly Asn Leu Pro Arg Glu Gly
432          515          520          525
435 Ala Val Leu Ile Val Thr Ala Ser Tyr Asn Gly His Pro Pro Asp Asn
436          530          535          540
439 Ala Lys Gln Phe Val Asp Trp Leu Asp Gln Ala Ser Ala Asp Glu Val
440 545          550          555          560
443 Lys Gly Val Arg Tyr Ser Val Phe Gly Cys Gly Asp Lys Asn Trp Ala
444          565          570          575
447 Thr Thr Tyr Gln Lys Val Pro Ala Phe Ile Asp Glu Thr Leu Ala Ala
448          580          585          590
451 Lys Gly Ala Glu Asn Ile Ala Asp Arg Gly Glu Ala Asp Ala Ser Asp
452          595          600          605
455 Asp Phe Glu Gly Thr Tyr Glu Glu Trp Arg Glu His Met Trp Ser Asp
456          610          615          620
459 Val Ala Ala Tyr Phe Asn Leu Asp Ile Glu Asn Ser Glu Asp Asn Lys
460 625          630          635          640
463 Ser Thr Leu Ser Leu Gln Phe Val Asp Ser Ala Ala Asp Met Pro Leu

```

VERIFICATION SUMMARY

DATE: 06/06/2002

PATENT APPLICATION: US/10/018,730A

TIME: 12:55:55

Input Set : A:\P02353US1.txt

Output Set: N:\CRF3\06062002\J018730A.raw

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date